



TITLE:

# 多変量解析論と情報科学の諸問題 (統計的多変量解析の研究報告集)

AUTHOR(S):

北川, 敏男

---

CITATION:

北川, 敏男. 多変量解析論と情報科学の諸問題 (統計的多変量解析の研究報告集). 数理解析研究所講究録 1968, 44: 133-144

ISSUE DATE:

1968-05

URL:

<http://hdl.handle.net/2433/107673>

RIGHT:

## 多変量解析論と情報科学の諸問題

九大理 数学教室 北川敏男

### I. 概要

0. 緒言. 標記の論題は、講演者にとっては統計学全般にわたる今後の発展方向がいかにあるべきかという問題を明らかにすることなしには立ち入って論ぜられないように思われる.

### 1. 序論

1.1 統計学の将来 The future of statistics (1967, 6月於 Wisconsin 大学)

1.2 統計学における計算機利用

(a) K. Pearson の時代 (b) R.A. Fisher - (F. Yates の会長就任講演)

(c) 現在の利用方式とその新しい方向

(i) TSS による共同利用

(ii) MAC による対話モード

> (バッチ方式との  
対比)

1.3 講演者の既往の論文

(a) (The recognition of statistics)

(1948)

(b) The logical aspects of successive processes of statistical inferences and controls, Bull. I.S.I., Vol. 38 (1961)

(c) The relativistic logic of mutual specification in statistics, Mem. Fac. Kyushu Univ., A 17 (1963)

(d) Automatically controlled sequence of statistical procedures in data analysis, Mem. Fac. Kyushu Univ., 17 (1963)

(e) Automatically controlled sequence of statistical procedures, Bernoulli-Baues-Laplace Anniversary Vol. IRS, Berkeley, (1965), Springer

(f) Information science and its connections with statistics, The 5th Berkeley Symposium (1965)

## 1.4. 例示と課題

例 1.1. Management Information System

例 1.2. Biomedical Computer Program

例 1.3. Medical Diagnosis System

- 課題 1. 情報源の多様化と情報処理の問題
- 課題 2. 巨大システムにおける情報と制御の問題
- 課題 3. パターン認識の問題
- 課題 4. 学習過程の問題

## 1.5. 本論文における 課題接近への4つの特徴

(1) 統計的認識の構造 (Structure of statistical recognition) (§2)

(2) 統計(的)処理のシステム解析 (System analysis of statistical procedures) (§3)

(3) 統計処理の統合方式 (Integration scheme of statistical procedures) (§4)

(4) 統計的認識における学習過程 (Learning processes in statistical recognition) (§5)

## 1.6. 多変量解析論との関連

## 2. Structure of statistical recognition

### 2.1. 情報処理と統計的認識の形成

#### (10) Information Processing

(形式) 伝達 — 変換 — 貯蔵

(意味) 認知 — 評価 — 指令

#### (20) 統計的認識の形成過程 (1a) 参照

(i) 集約化 (ii) 標識化 (iii) 差別化

(iv) 数値化 (v) 等質化 (vi) 不確率化

#### (30) 統計的認識と制御

(i) 因子の導入 — その種別 (実験計画法 形式: 参照)

(ii) オペレーションと機構 (形式: 参照)

#### (40) 統計的認識とその用途 (1c) 参照

(i) Storage (貯蔵) ... S

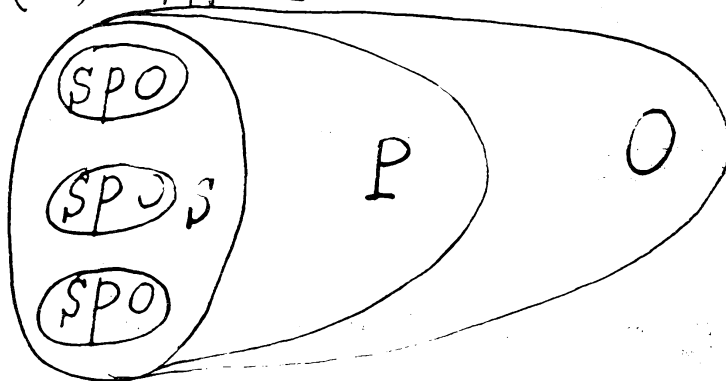
(ii) Pattern Recognition (パターン認識) ... P

(iii) Operational uses (オペレーション用) ... O

### 2.2. 統計的認識の構造

#### (10) 階層的構造 (heirarchical structure)

#### (20) 構造の変化, 再編成



$S \rightarrow P \rightarrow O$

## 2.3. 不完全規定模型とパターン認識

例2.1. Pooling of data TE型推測

例2.2. 適応制御における同定問題

$$Ay(t) = x(t), \quad x(0) = C$$

(C, x(t), y(t)) 既知のとき A? を推定する.

基本的な視覚: パターン認識と operation との関係にある. estimation problem との相違点

## 2.4. 統計的認識と統計的処理

— Postulation of statistical procedures —

(1°) (i) 可能な前提群の幾組  $\alpha_1, \alpha_2, \dots, \alpha_m$  をつくる.(ii) 可能な評価基準群の幾組  $C_1, C_2, \dots, C_n$  をつくる.(iii) 前提群  $\times$  評価基準群の組合せ  $(\alpha_i, C_j)$  ( $i=1 \sim m; j=1 \sim n$ ) の各々に対して 解  $S_{ij}$  を求める.(2°)  $\{S_{ij}\}$  ( $i=1, 2, \dots, m; j=1, 2, \dots, n$ ) のなかの選択様態は, 或はそのなかの“最適なもの”は, 他の information との結合によって判断する.  
(あるいは)

(3°) 以上の方法において, interpenetrating principle が用いられる. これは mutual specification の論理とつながる. (論文(c)参照)

例2.3

{	母集団	{	正規型	$\alpha_1$
			type II 型	$\alpha_2$
			対数正規型	$\alpha_3$
	最尤法 — 最小分散			
	$C_1$		$C_2$	

### 3. System analysis of statistical procedures

#### 3.1. Cybernetical formulation

サイバネティクス の観点からのシステム構成

- (i) 入力 → 暗箱 → 出力
- (ii) 暗箱の連結
- (iii) フィード・バック回路の存在
- (iv) 暗箱の内部分解
- (v) 人間・機械系の立場 (全体構成)

#### 3.2. Statistical programming

統計処理のためのコンピューターのプログラミング

(g) Kendall, M. G., and P. Wegner: An Introduction to Statistical programming. ISI (1965) Ottawa

(h) Terry, M. E.: The principles of statistical analysis using large electronic computers ISI (1965) Ottawa

(i) Yates, F., and H. R. Simpson: The analysis of surveys: The Computer Journal 4 (1961)

(10) Terry の 4 原則とその意味 (論文 (c) § 4 号四)

(20) 推測過程論との関連

(30) conversational mood との関係

(40) MAC と ACSSP との関係

## 3.3 システム解析的接近

## System analysis approach

## (1) 構成関係

成分分析	component analysis
結合分析	connection analysis
構成分析	composition analysis

## (2) 変動関係

安定性分析	stability analysis
信頼性分析	reliability analysis
融通性分析	flexibility analysis

	入力	システム	出力
安定	○		
信頼		○	
融通			○

## 3.4 統計処理の系列 (Sequence of statistical procedures)

(10) 個々の道

(20) 可能な道の集合 — 樹木の概念

(30) 確率概念の導入の可能性

## 4. 統計的認識の統合方式 Integration Scheme of Statistical Recognition

### 4.1. 統合方式 Integration Scheme

(1) 情報集約方式 Information Summery Scheme

(2) 情報評価方式 Information Evaluation Scheme

(論文 (c) 参照)

### 4.2. 統合方式の2つの視点

(1) 機能 (伝達 - 変換 - 蓄積)

(2) 構成 (成分 - 結合 - 構成)

(3) 変動 (安定 - 信頼 - 融通)

(4) 意味 (認知 - 評価 - 指令)

### 4.3. 例示

例4.1. estimation problem

- (i) operational use
- (ii) storage
- (iii) pattern recognition

例4.2. testing hypothesis

- ハロウ - タ - 空間
- fuzzy set



## 5. 統計的認識における学習過程

### 5.1. 統合方式の可変性

型	集約方式	評価方式
(1)	不変	不変
(2)	可変	不変
(3)	不変	可変
(4)	可変	可変

### 5.2. 諸理論からの例示

例 5.1.	マルコフ型学習過程	(1)
例 5.2.	Iterative operation	(1)
例 5.3.	逐次解析	(1)
例 5.4.	推測過程	(2), (3), (4)
例 5.5.	Learntol	(3)
例 5.6.	ACSSP	(2), (3), (4)

### 5.3. 学習過程の必要

#### 例 5.7. 医療システム

### 5.4. Heuristics

(10) conversational mood への learning process

(20) heuristics の方法

(理論) — (実験) — (方式)

## II 補註

概要は、簡潔に要旨を箇条書きにまとめたものである。  
以下、これを補うために次の補註を加える。

1.1. 「統計学の将来」というシンポジウムが G. E. Box, G. Barnard, J. Tukey, O. Kempthorne 等によって行われてることを紹介した。筆者も招待をうけたが、出席できなかった。

1.2. これは計算機利用の新しい方式がひたおしする影響を指摘するのが目的である。

2.12 R. A. Fisher がハフチンズ変換式に不満を述べた講話等が紹介される。従って (C) (i), (ii) が見逃し難いことを指摘する。

1.3. 2.11 は (a) から (f) までの筆者の論文の振替を紹介している。2の論文の主旨は、さらに進展しなければならぬことを強調する点にある。

1.4. 1.3の趣意を3つの例をあげて具体的に説明すると共に、2.11のあいさとは4つの課題をあげている。2.11によつて情報子的接近の要請をいふことを明示するのが趣意である。

1.5. 上述のような課題に対するこれにいかん攻策するか、基本的には4つの接近路がある。

1.6. これは29シンボリックが多変量解析論をテーマにした図像もあつて、この項で改めた。筆者の積極的<sup>に</sup>主張したいことは、多変量解析とは、本稿のような視座からの接近が必要にある。それにもかかわらず、安易に一変量解析の類似問題<sup>に</sup>あまり区別がつかないのではなかろうかということである。

2.1. これは拙著「統計学の認識」と拙著「実験計画法講義」との2つの見方を統合し整理したものがある。S (storage), P (pattern recognition), O (operational uses) という三面をいいあげているのは拙論文(C)に沿っている。

2.2. S, P, Oを要素としての構造説明がある。2のような構造については、数多くの統計的認識を広く検討するとか大切なことがある。現代の数値統計学は、2の意味、統計的認識の諸形態をくまなく探索すること、充分力を入れている。

2.3 及び 2.4 は 2.2 をより具体的に述べて積りである。

3.1 cybernetical formulation—これは

については、すでに論文 (b) において觸れている。しかしシステムという概念の導入は、(c) においてあった。その予備的考察とあるものの、3.1 の目的がある。

3.2. Statistical programming という用語については詳しく述べるのは M. G. Kendall 達である。筆者の論文 (e) では、(g) - (i) について詳しく言及している。その報告は (d) において新しい面も指摘している。それは (3) の (f) である。

3.3 は 3.1 をより具体的に述べるものである、そのあとの論述は、さう一般と具体化し分けなければならない。

4.1 の integration scheme, 4.2 の視座いずれも、情報論的な視野を展開しないわけにはゆかない。そしてその見地から 4.3 の例示に入る。例えば 'estimation problem', 例えば 'testing hypothesis' 24) の古典的問題とも見らなければならない (情報論的な見地からすれば)、多くの得点がその型式化のふたつの点にある。fuzzy set (制御工学者 Zadeh 達) などが、2.1.1 にも注意されたい。

5. においては 学習過程を伴う 統計的領域で問題にする。2.4.6 統合方式の可変性において、  
 うえようとする 筆者の立場では、型(1), (2), (3), (4)が 当分 あり得る。2.4.7 5.1 において述べている。

6.2 には この例示を、筆者自身の研究例、その他の例が 集めて、より 具体的にする。

6.3 には 学習過程の 必要なることを 医療システムに述べている。

6.4 には *heuristic* の 追々進んで述べて、2の項は 詳しく述べて 2と1の 差を挙げる。

以上が この論文の 趣意である。

筆者の考えは、(a)から (f)まで 進んでいくが、しかし 2.1 において、一般と 改変し、その *flexibility* を 増加させる 必要がある。2.4.6 また 2.5 には 一歩一歩 進む 2.1 といふわけでも、いままで 2.1 があったように 筆者自身は 2.1 として 進んでいく と思う。シンボリックであるから、自由な 意見を述べて 2.1 といふ。しかし 実質的な内容 づけをする仕事は 2.1 に見よう、2.1 2.2 の 2.1 といふ。(1958年3月8日記)